Optimal Binary Locally Repairable Codes via Anticodes

Natalia Silberstein and Alexander Zeh
Computer Science Department
Technion—Israel Institute of Technology
Haifa 32000, Israel
natalys@cs.technion.ac.il, alex@codingtheory.eu

Abstract—**This paper presents a construction for several families of optimal binary locally repairable codes (LRCs) with small locality (2 and 3). This construction is based on various anticodes. It provides binary LRCs which attain the Cadambe–Mazumdar bound. Moreover, most of these codes are optimal with respect to the Griesmer bound.**

I. INTRODUCTION

Locally repairable codes (LRCs) are a family of erasure codes which allow local correction of erasures, where any code symbol can be recovered by using a small (fixed) number of other code symbols. The concept of LRCs was motivated by application to distributed storage systems (DSSs), (see e.g. [2], [3] and references therein). DSSs store data across a network of nodes in a redundant form to ensure resilience against node failures. Using of LRCs to store data in DSSs enables to repair a failed node locally, i.e., by accessing a small number of other nodes in the system.

The ith code symbol c_i, $1 \le i \le n$, of an $[n, k, d]$ linear code $\mathcal{C}$ is said to have *locality* r if c_i can be recovered by accessing at most r other code symbols. A code $\mathcal{C}$ is said to have locality r if all its symbols have locality r. Such codes are referred to as locally repairable (or recoverable) codes (LRCs). LRCs were introduced by Gopalan *et al.* in [6]. It was shown in [6] that the minimum distance of an $[n, k, d]$ LRC with locality r should satisfy the following generalization of the Singleton bound

$$d \le n - k + 2 - \left\lceil \frac{k}{r} \right\rceil. \tag{1}$$

Constructions of LRCs which attain this bound were presented in [5], [6], [20], [21]. Further generalizations of LRCs to the codes which can locally correct more than one erasure, to the LRCs with multiple repair alternatives, and to the vector LRCs were considered in [8], [9], [11]–[16], [18], [20].

However, to attain the bound (1) and its generalizations [8], [12], [13], [16], the known codes should be defined over a large finite field. In [20] Tamo and Barg presented LRCs satisfying the bound (1), which are defined over a field of size slightly greater than the length of the code. This is the smallest field size for optimal LRCs known so far.

Codes over small (especially binary) alphabets are of a particular interest due to their implementation ease. Recently,

a new bound for LRCs which takes the size of the alphabet into account was established by Cadambe and Mazumdar [1, Thm. 1]. They showed that the dimension k of an $[n, k, d]$ LRC $\mathcal{C}$ over $\mathbb{F}_q$ with locality r is upper bounded by

$$k \le \min_{t \in \mathbb{Z}^+} \left\{ tr + k_{opt}^{(q)} \big(n - t(r+1), d \big) \right\}, \tag{2}$$

where $k_{opt}^{(q)}(n, d)$ is the largest possible dimension of a code of length n, for a given alphabet size q and a given minimum distance d. This bound applies to both linear and nonlinear codes. In the case of a nonlinear code, the parameter k is defined as $|\mathcal{C}| / \log q$. Moreover, it was shown in [1] that the family of binary simplex codes attains the bound (2) for $r = 2$. To the best of our knowledge, there are only three additional works that consider constructions of binary LRCs, namely the papers of Shahabinejad *et al.* [19], Goparaju and Calderbank [7], and Zeh and Yaakobi [22], where the constructions of [7] and [22] consider binary cyclic LRCs.

In this paper we propose constructions of new binary LRCs which attain the bound (2). All our LRCs have a small locality ($r = 2$ and $r = 3$), moreover, most of our codes attain the Griesmer bound. Our constructions use a method of Farrell [4] based on anticodes. In particular, we modify a binary simplex code by deleting certain columns from its generator matrix. These deleted columns form an anticode. We investigate the properties of anticodes which allow constructions of LRCs with small locality. Also, we present optimal binary LRCs with locality 2 based on subspace codes.

The rest of the paper is organized as follows. In Section II we provide the necessary definitions, in particular, we define anticodes and describe a method of constructing a new code based on a simplex code and an anticode. In Section III we present our constructions based on various choices of anticodes. Conclusion is given in Section IV.

II. PRELIMINARIES

Let $\mathcal{C}$ be a *linear* $[n, k, d]$ *code* of length n, dimension k and minimum Hamming distance d over $\mathbb{F}_q$. We say that a $k \times n$ generator matrix G of $\mathcal{C}$ is in a *standard form* if it has the form $G = (I_k | A)$, where I_j is an identity matrix of order j and A is a $k \times (n - k)$ matrix. If G is in a standard form then an $(n-k) \times n$ parity-check matrix H can be easily obtained from G in the following way: $H = (-A^T | I_{n-k})$. Note that for a binary code, we have $H = (A^T | I_{n-k})$. In the

sequel, we will consider only binary codes. The following simple lemma shows a sufficient condition on a parity-check matrix of a code with locality r.

Lemma 1. *An $[n, k, d]$ linear code has locality r if for every coordinate i, $1 \leq i \leq n$, there exists a row R_i of weight at most $r + 1$ in its parity-check matrix, which has a nonzero entry in the ith coordinate. In this case we say that the coordinate i is covered by the row R_i.*

The following two bounds will be used in the sequel. The Plotkin bound (Thm. 1) holds for nonlinear codes, while the Griesmer bound is restricted to linear codes (Thm. 2).

Theorem 1 (Plotkin Bound [10, p. 43]). *Let $A_2(n, d)$ denote the largest number of codewords in a binary code of length n and minimum distance d. If d is even and $2d > n$ then*

$$A_2(n, d) \leq 2 \left\lfloor \frac{d}{2d - n} \right\rfloor.$$

Theorem 2 (Griesmer Bound [10, p. 547]). *The length n of a binary linear code with dimension k and minimum distance d must satisfy*

$$n \geq \sum_{i=0}^{k-1} \left\lceil \frac{d}{2^i} \right\rceil.$$

In the remaining part of this section we define an anticode and recall the anticode-based construction of binary linear codes by Farrell [4] (see also [10, p. 548]). An *anticode* is a code which may contain repeated codewords and which has an *upper* bound on the distance between the codewords. More precisely, a binary linear anticode $\mathcal{A}$ of length n and maximum distance δ is a set of codewords in $\mathbb{F}_2^n$ such that the Hamming distance between any pair of codewords is less than or equal to δ. The generator matrix $G_{\mathcal{A}}$ of $\mathcal{A}$ is a $k \times n$ binary matrix such that all the 2^k combinations of its rows form the codewords of the anticode. If $\text{rank}(G_A) = \gamma$, then each codeword occurs $2^{k-\gamma}$ times in the anticode. Due to linearity, we have

$$\delta = \max_{\mathbf{a} \in \mathcal{A}} \text{wt}(\mathbf{a}), \tag{3}$$

where $\text{wt}(\mathbf{v})$ denotes the Hamming weight of a vector $\mathbf{v}$.

Example 1. *Let $G_{\mathcal{A}}$ be a 3×3 generator matrix given by*

$$G_{\mathcal{A}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

It generates a binary linear anticode $\mathcal{A}$ of length $n = 3$ and $\delta = 2$, where the set of 2^3 codewords is

$$\mathcal{A} = \big\{ (000), (110), (101), (011), (011), (101), (110), (000) \big\}.$$

The construction of Farrell [4] which we use to construct optimal LRCs is based on a modification of a generator matrix for a binary simplex code.

A binary *simplex* code $\mathcal{S}_m$ is a $[2^m - 1, m, 2^{m-1}]$ code with generator matrix G_m whose columns consist of all distinct nonzero vectors in $\mathbb{F}_2^m$. In the rest of the paper we assume w.l.o.g. that G_m is in the standard form and that the columns of G_m are ordered according to their Hamming weight.

Construction 1 (Farrell Construction [4]). *Let G_m be the $m \times (2^m - 1)$ generator matrix of a binary simplex code $\mathcal{S}_m$ and let $G_{\mathcal{A}}$ be the $k \times n$ generator matrix with distinct columns of a binary linear anticode $\mathcal{A}$ of length n and maximum distance δ. Then, the $m \times (2^m - 1 - n)$ matrix obtained by deleting the n columns of $G_{\mathcal{A}}$ from G_m is a generator matrix of a binary $[2^m - 1 - n, \leq m, 2^{m-1} - \delta]$ code.*

Example 2. *Let G_4 be the 4×15 generator matrix of a simplex code $\mathcal{S}_4$ and let $G_{\mathcal{A}}$ be the generator matrix of the anticode given in Example 1 with the additional first row of zeros. By deleting the columns of $G_{\mathcal{A}}$ from G_4 we obtain the following matrix*

$$G_4 \setminus G_{\mathcal{A}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

where the shadowed columns in $\{8, 9, 10\}$ are deleted, which generates a $[12, 4, 6]$ code. Note that this code attains the Griesmer bound.

III. CONSTRUCTIONS OF OPTIMAL BINARY LRCS

In this section we provide constructions of binary LRCs based on the Farrell construction (see Construction 1), by using various anticodes. We prove that our codes have a small locality ($r = 2$ or $r = 3$) and attain the Cadambe–Mazumdar bound (2). Most of our codes also attain the Griesmer bound (see Thm. 2).

A. LRCs based on Anticodes

First, we generalize Example 1 and consider an anticode such that all the vectors of length s and weight 2 form the columns of its generator matrix. We denote such an anticode by $\mathcal{A}_{s,2}$. For example, the anticode from Example 1 is an $\mathcal{A}_{3,2}$ anticode. First, we need the following theorem about the parameters of $\mathcal{A}_{s,2}$.

Theorem 3. *Let $\mathcal{A}_{s,2}$ be a binary anticode such that all weight-2 vectors of length s form the columns of its generator matrix $G_{\mathcal{A}}$. Then $\mathcal{A}_{s,2}$ has length $\binom{s}{2}$ and maximum weight $\delta = \left\lfloor \frac{s^2}{4} \right\rfloor$.*

Proof: There are $\binom{s}{2}$ vectors of weight 2 and length s, hence the number of columns in $G_{\mathcal{A}}$ and the length of $\mathcal{A}_{s,2}$ is $\binom{s}{2}$. Next, we prove that the value of the maximum weight δ is $\left\lfloor \frac{s^2}{4} \right\rfloor$. Note that the $s \times \binom{s}{2}$ generator matrix $G_{\mathcal{A}}$ is also an incidence matrix of a complete graph $K_s = (V, E)$ with $|V| = s$ and $|E| = \binom{s}{2}$. Therefore, the maximum weight δ of the anticode can be described in terms of maximum cut between a subset of vertices $S \subseteq V$ and its complement S^c, more precisely,

$$\delta = \max_{1 \leq i \leq s} |\text{Cut}(S_i, S_i^c)|,$$

where S_i is a subset of V of size i and S_i^c its complement of size $s - i$. Since K_s is an $(s - 1)$-regular graph, it holds that for all $1 \leq i \leq s$, $(s - 1)i = |\text{Cut}(S_i, S_i^c)| + 2|E_i|$, where $E_i \subseteq E$ is the set of edges between the vertices in S_i. Note

that the induced subgraph (S_i, E_i) of K_s is a complete graph K_i and then $|E_i| = \binom{i}{2}$. Thus,

$$|\text{Cut}(S_i, S_i^c)| = (s-1)i - i(i-1) = i(s-i)$$

and

$$\delta = \max_{1 \leq i \leq s} \{i(s-i)\} = \left\lfloor \frac{s^2}{4} \right\rfloor.$$

∎

As a consequence of Thm. 3 and the Farrell construction we have the following theorem.

Theorem 4. *Let $\mathcal{S}_m$ be a $[2^m - 1, m, 2^{m-1}]$ simplex code, $m \geq 4$, and let G_m be its generator matrix. Let $\mathcal{A}_{s,2}$, $s \leq m$, be an anticode defined in Thm. 3 and let G_A be its generator matrix. We prepend $m - s$ zeros to every column of G_A to form an $m \times \binom{s}{2}$ matrix whose columns are deleted from G_m to obtain a generator matrix G for a new code $\mathcal{C}_{m,s,2}$. Then $\mathcal{C}_{m,s,2}$ is a $[2^m - \binom{s}{2} - 1, m, 2^{m-1} - \lfloor \frac{s^2}{4} \rfloor]$ LRC with locality $r = 2$.*

Proof: The prepending of zeros to G_A does not change the length and the maximum weight of the anticode. Hence, the length, the dimension, and the minimum distance of the obtained code $\mathcal{C}_{m,s,2}$ directly follow from the Farrell construction and Thm. 3.

To prove that the locality of $\mathcal{C}_{m,s,2}$ is $r = 2$, by Lemma 1 we need to show that every coordinate is covered by a row of weight 3 of the parity-check matrix for $\mathcal{C}_{m,s,2}$ (note that clearly locality is not 1). Since G_m is in the standard form, the generator matrix G of $\mathcal{C}_{m,s,2}$ is also in the standard form. We denote by G^i, $1 \leq i \leq m$, the submatrix of G which consists of the set of columns of G of weight i. Then the parity-check matrix H of $\mathcal{C}_{m,s,2}$ has the following form:

$$H = \left(\begin{array}{c|c|c|c|c} (G^2)^T & I_{\binom{m}{2}-\binom{s}{2}} & & & \\ \hline \vdots & & \ddots & & \\ \hline (G^{m-1})^T & & & I_m & \\ \hline (G^m)^T & & & & 1 \end{array} \right).$$

We will show that by a simple modification of H with elementary operations on its rows we obtain a parity-check matrix H' such that every coordinate of the code will be covered by a row of H' of weight 3. We define a partition of columns of H into the parts $\{H^1, \ldots, H^m\}$ as follows. The part H^1 contains the first m columns, and the part H^i, $2 \leq i \leq m$, corresponds to the columns which contain $I_{\binom{m}{i}}$ included in the rows which contain $(G^i)^T$. (Note that if $s = m$ then $H^2 = \emptyset$.) Let consider the xth coordinate in H^i, $3 \leq i \leq m-1$. There exists a row R_x^i in H with nonzero entry in this coordinate. This row R_x^i also contains i nonzero entries in the first m coordinates. Let R_y^{i+1} be a row in H such that its $i+1$ nonzero entries in the first m coordinates contain the first i nonzero entries of R_x^i, and which also has one in the yth coordinate of H^{i+1}. Then the coordinates x in H^i and y in H^{i+1} are covered by $R_x^i + R_y^{i+1}$, the weight-3 row of H'. Note that for every x in H^i there is such y in H^{i+1}. To show that any coordinate in $H^1 \cup H^m$ has locality 2, note that when we add the last row of H to each one of the m rows which contain the rows of $(G^{m-1})^T$,

we obtain m rows of weight 3 in H' with the first nonzero entry in the first m coordinates and the last nonzero entry in the last coordinate. Thus, the obtained parity-check matrix H' contains weight-3 rows which cover all the coordinates (and the last row of H, of weight $m + 1$).

∎

In the following example we illustrate the idea of modification of a parity-check matrix described in the proof of Thm. 4.

Example 3. *We consider the parity-check matrix H of the $[12, 4, 6]$ code of Example 2 based on the anticode $\mathcal{A}_{3,2}$. It has the following form, where the vertical lines show the partition of its columns into the parts H^1, H^2, H^3, H^4.*

$$H = \left(\begin{array}{cccc|ccc|cccc|c} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

We add to each one of the ith rows of H, $4 \leq i \leq 7$, the last row and obtain the following parity-check matrix:

$$H' = \left(\begin{array}{cccc|ccc|cccc|c} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right),$$

such that every coordinate is covered by a weight-3 row.

Corollary 1. *For $3 \leq s \leq 5$, the code $\mathcal{C}_{m,s,2}$ obtained in Thm. 4 attains the bound (2). More precisely,*

- *The code $\mathcal{C}_{m,3,2}$ obtained by using the anticode $\mathcal{A}_{3,2}$ is a $[2^m - 4, m, 2^{m-1} - 2]$ LRC with locality $r = 2$ which attains the bound (2).*
- *The code $\mathcal{C}_{m,4,2}$ obtained by using the anticode $\mathcal{A}_{4,2}$ is a $[2^m - 7, m, 2^{m-1} - 4]$ LRC with locality $r = 2$ which attains the bound (2).*
- *The code $\mathcal{C}_{m,5,2}$ obtained by using the anticode $\mathcal{A}_{5,2}$ is a $[2^m - 11, m, 2^{m-1} - 6]$ LRC with locality $r = 2$ which attains the bound (2).*

Proof: To prove the optimality of the proposed codes we apply the bound (2) with $t = 1$ and use the Plotkin bound (see Thm. 1):

For $s = 3$ we have $2 + k_{opt}^{(2)}(2^m - 7, 2^{m-1} - 2) \leq 2 + \left\lfloor \log 2 \left\lfloor \frac{2^{m-1}-2}{3} \right\rfloor \right\rfloor \leq 2 + \lfloor \log(2^{m-1} - 2) \rfloor = 2 + m - 2 = m$.

For $s = 4$ we have $2 + k_{opt}^{(2)}(2^m - 10, 2^{m-1} - 4) \leq 2 + \left\lfloor \log 2 \left\lfloor \frac{2^{m-1}-4}{2} \right\rfloor \right\rfloor = 2 + \lfloor \log(2^{m-1} - 4) \rfloor = 2 + m - 2 = m$.

For $s = 5$ we have $2 + k_{opt}^{(2)}(2^m - 14, 2^{m-1} - 6) \leq 2 + \left\lfloor \log 2 \left\lfloor \frac{2^{m-1}-6}{2} \right\rfloor \right\rfloor = 2 + \lfloor \log(2^{m-1} - 6) \rfloor = 2 + m - 2 = m$.

∎

Remark 1. *One can check that the codes $\mathcal{C}_{m,3,2}$ and $\mathcal{C}_{m,5,2}$ attain the Griesmer bound.*

Note that to apply our modification of a parity-check matrix in the proof for locality 2, the generator matrix of

a code obtained by the Farrell construction should contain columns of consecutive weights. Based on this observation, we propose a generalization of the previous construction of an anticode and prove that the LRCs obtained from this anticode have locality $r = 2$ and attain the Griesmer bound.

Theorem 5. *Let $\mathcal{S}_m$ be a $[2^m - 1, m, 2^{m-1}]$ simplex code, $m \geq 4$, and let G_m be its generator matrix. Let $\mathcal{A}_{t;2,3,\ldots,t-1}$, $3 \leq t \leq m$, be an anticode such that its generator matrix $G_{\mathcal{A}}$ consists of all the columns in $\mathbb{F}_2^t$ of weights in $\{2, 3, \ldots, t - 1\}$. We prepend $m - t$ zeros to every column of $G_{\mathcal{A}}$ to form the $m \times \sum_{i=2}^{t-1} \binom{t}{i}$ matrix whose columns are deleted from G_m to obtain a generator matrix G for a new code $\mathcal{C}_{m,t}$. Then $\mathcal{C}_{m,t}$ is a $[2^m - 2^t + t + 1, m, 2^{m-1} - 2^{t-1} + 2]$ LRC with locality $r = 2$ which attains the Griesmer bound.*

Proof: First we prove that the anticode $\mathcal{A}_{t;2,3,\ldots,t-1}$ has length $2^t - t - 2$ and maximum weight $2^{t-1} - 2$. Note that the generator matrix of $\mathcal{A}_{t;2,3,\ldots,t-1}$ can be obtained from the generator matrix G_t of the simplex code $\mathcal{S}_t$ by removing t columns of weight 1 and one column of weight t, and hence the length of $\mathcal{A}_{t;2,3,\ldots,t-1}$ is $2^t - t - 2$. Since all the codewords in $\mathcal{S}_t$ have weight 2^{t-1} and from each row of G_t we removed two ones to obtain a generator matrix for our anticode, where all the rows in G_t have one of the removed ones in the same place, it follows that the maximum weight of $\mathcal{A}_{t;2,3,\ldots,t-1}$ is $\delta = 2^{t-1} - 2$. Hence, since prepending zero rows to $G_{\mathcal{A}}$ does not change the length and the maximum weight of the anticode, $\mathcal{C}_{m,t}$ is a $[2^m - 2^t + t + 1, m, 2^{m-1} - 2^{t-1} + 2]$ code. The proof of locality is similar to the proof in Thm. 3.

To prove that $\mathcal{C}_{m,t}$ attains the Griesmer bound we have

$$\sum_{i=0}^{m-1} \left\lceil \frac{2^{m-1} - 2^{t-1} + 2}{2^i} \right\rceil = \sum_{i=0}^{m-1} 2^i - \sum_{i=0}^{t-1} 2^i + (2 + t - 1)$$

$$= 2^m - 1 - 2^t + 1 + t + 1 = 2^m - 2^t + t + 1,$$

which completes the proof. ∎

In the following, we consider an anticode formed by another modification of a simplex code and prove that the LRC obtained from this anticode attains the bound (2) and has locality 3.

Theorem 6. *Let $\mathcal{A}_{m-1}$ be the anticode with generator matrix $G_{\mathcal{A}}$ given by*

$$G_{\mathcal{A}} = \left(\begin{array}{c|c} 1 & 000\ldots00 \\ \hline 0 & \\ \vdots & G_{m-1} \\ 0 & \end{array} \right),$$

where G_{m-1} is the generator matrix for the simplex code $\mathcal{S}_{m-1}$. Let $\mathcal{C}$ be a code obtained by the Farrell construction based on the simplex code $\mathcal{S}_m$ and on the anticode $\mathcal{A}_{m-1}$. Then $\mathcal{C}$ is a $[2^{m-1} - 1, m, 2^{m-2} - 1]$ LRC with locality $r = 3$ which attains the bound (2).

Proof: Since all the codewords in $\mathcal{S}_{m-1}$ have the constant weight 2^{m-2}, the maximum weight of $\mathcal{A}_{m-1}$ is $\delta = 2^{m-2} + 1$. Then, by the Farrell construction, $\mathcal{C}$ is a $[2^m - 1 - 2^{m-1}, m, 2^{m-1} - 2^{m-2} - 1] = [2^{m-1} - 1, m, 2^{m-2} - 1]$

code. Note that $\mathcal{C}$ is the augmented simplex code $\mathcal{S}_{m-1}$ with generator matrix:

$$G = \left(\begin{array}{c} 111\ldots11 \\ \hline G_{m-1} \end{array} \right).$$

To prove that the locality is 3, we note that all the codewords in the dual code of $\mathcal{C}$ have even weight, and hence the locality r is an odd number. Clearly, $r > 1$. We construct the parity-check matrix H of $\mathcal{C}$ with all the rows of weight 4, and then every coordinate will be covered by a row of H of weight 4, which by Lemma 1 implies that $r = 3$. Recall that the dual code of $\mathcal{S}_{m-1}$ is a $[2^{m-1}, 2^{m-1} - m, 3]$ Hamming code [10], and denote its generator matrix by H^{m-1}. We consider the construction of H^{m-1} with all the rows of weight 3 from [1], where the rows have nonzero entries in the positions $(i, 2^j, i + 2^j)$, for $1 \leq j \leq m - 2$, $1 \leq i \leq 2^j$. Let denote by $H_{i,2^j}^{m-1}$ a row of H^{m-1} with nonzero entries in the positions $(i, 2^j, i + 2^j)$. The parity-check matrix H will consist of $2^{m-1} - m - 1$ weight-4 rows $H_{1,2}^{m-1} + H_{1,2^j}^{m-1}$, $2 \leq j \leq m - 2$, and $H_{i,2^j}^{m-1} + H_{i+1,2^j}^{m-1}$, $2 \leq j \leq m - 2$, $1 \leq i \leq 2^j - 2$.

To prove the optimality of the obtained code $\mathcal{C}$ we apply the bound (2) with $t = 1$ and use the Plotkin bound:

$$3 + k_{opt}^{(2)}(2^{m-1} - 1 - 4, 2^{m-2} - 1) \leq 3 + \left\lfloor \log 2 \cdot \left\lfloor \frac{2^{m-2} - 1}{3} \right\rfloor \right\rfloor$$

$$\leq 3 + \left\lfloor \log(2^{m-2} - 1) \right\rfloor = 3 + m - 3 = m.$$

∎

Remark 2. *One can check that the code $\mathcal{C}$ from Thm. 5 attains the Griesmer bound.*

B. LRCs based on Subspace Codes

In this subsection we consider a construction of optimal binary LRCs with the parity-check matrix formed by the codewords of a *subspace* code. In particular, we are interested in a special kind of subspace codes, called *lifted rank-metric codes* [17], with the trivial distance 1 and the constant dimension 2. More precisely, let consider a $2^{2s-4} \times (2^s - 2^{s-2})$ binary matrix H^s whose columns are indexed by the vectors in $\mathbb{F}_2^s \setminus \{00v : v \in \mathbb{F}_2^{s-2}\}$ and whose rows are indexed by the 2-dimensional subspaces of $\mathbb{F}_2^s$ contained in $\mathbb{F}_2^s \setminus \{00v : v \in \mathbb{F}_2^{s-2}\}$. Every row of H^s is the incidence vector of such a 2-dimensional subspace and then has weight 3 (note that the all-zero vector is not considered). It was proved in [17, Thm. 11] that the code $\mathcal{C}^s$ with the parity-check matrix H^s is a $[2^s - 2^{s-2}, s, \frac{2^s - 2^{s-2}}{2}]$ 2^{s-2}-quasi-cyclic code. Note that H^s contains dependent rows.

Example 4. *For $s = 4$, the matrix H^4 has the following form. The four rows above this matrix represent the vectors which index the columns of H^4. For example, the first row of H^4 corresponds to the subspace which contains the vectors $\{(0, 1, 0, 0), (1, 0, 0, 0), (1, 1, 0, 0)\}$:*

$$\begin{array}{ccc}
\mathbf{0\ 0\ 0\ 0} & \mathbf{1\ 1\ 1\ 1} & \mathbf{1\ 1\ 1\ 1} \\
\mathbf{1\ 1\ 1\ 1} & \mathbf{0\ 0\ 0\ 0} & \mathbf{1\ 1\ 1\ 1} \\
\mathbf{0\ 0\ 1\ 1} & \mathbf{0\ 0\ 1\ 1} & \mathbf{0\ 0\ 1\ 1} \\
\mathbf{0\ 1\ 0\ 1} & \mathbf{0\ 1\ 0\ 1} & \mathbf{0\ 1\ 0\ 1}
\end{array}$$

$$\left(\begin{array}{ccc|ccc|ccc}
1\ 0\ 0\ 0 & 1\ 0\ 0\ 0 & 1\ 0\ 0\ 0 \\
0\ 1\ 0\ 0 & 0\ 0\ 0\ 1 & 0\ 0\ 1\ 0 \\
0\ 0\ 1\ 0 & 0\ 1\ 0\ 0 & 0\ 0\ 0\ 1 \\
0\ 0\ 0\ 1 & 0\ 0\ 1\ 0 & 0\ 1\ 0\ 0 \\
\hline
1\ 0\ 0\ 0 & 0\ 1\ 0\ 0 & 0\ 1\ 0\ 0 \\
0\ 1\ 0\ 0 & 0\ 0\ 1\ 0 & 0\ 0\ 0\ 1 \\
0\ 0\ 1\ 0 & 1\ 0\ 0\ 0 & 0\ 0\ 1\ 0 \\
0\ 0\ 0\ 1 & 0\ 0\ 0\ 1 & 1\ 0\ 0\ 0 \\
\hline
1\ 0\ 0\ 0 & 0\ 0\ 1\ 0 & 0\ 0\ 1\ 0 \\
0\ 1\ 0\ 0 & 0\ 1\ 0\ 0 & 1\ 0\ 0\ 0 \\
0\ 0\ 1\ 0 & 0\ 0\ 0\ 1 & 0\ 1\ 0\ 0 \\
0\ 0\ 0\ 1 & 1\ 0\ 0\ 0 & 0\ 0\ 0\ 1 \\
\hline
1\ 0\ 0\ 0 & 0\ 0\ 0\ 1 & 0\ 0\ 0\ 1 \\
0\ 1\ 0\ 0 & 1\ 0\ 0\ 0 & 0\ 0\ 1\ 0 \\
0\ 0\ 1\ 0 & 0\ 0\ 1\ 0 & 1\ 0\ 0\ 0 \\
0\ 0\ 0\ 1 & 0\ 1\ 0\ 0 & 0\ 0\ 1\ 0
\end{array}\right)$$

In the following we prove that this code attains the bound (2) with locality 2.

Theorem 7. *Let $\mathcal{C}^s$ be the linear code with the parity-check matrix H^s defined above. Then $\mathcal{C}^s$ is a $[3 \cdot 2^{s-2}, s, 3 \cdot 2^{s-3}]$ LRC with locality $r = 2$ which attains the bound (2).*

Proof: The length, the dimension and the minimum distance are proved in [17, Thm. 11]. Since every row in H^s has weight 3, the locality is 2. By applying the bound (2) with $t = 1$ and using the Plotkin bound we have

$$2 + k_{opt}^{(2)}(3 \cdot 2^{s-2} - 3, 3 \cdot 2^{s-3}) \le 2 + \log 2 \cdot \left\lfloor \frac{3 \cdot 2^{s-3}}{3} \right\rfloor$$

$$= 2 + \log 2^{s-2} = 2 + s - 2 = s$$

In other words, the code $\mathcal{C}^s$ always attain the bound (2). ∎

Remark 3. *One can check that the code $\mathcal{C}^s$ from Thm. 7 attains the Griesmer bound.*

Remark 4. *Note that the code $\mathcal{C}^s$ from Thm. 7 can be also constructed by applying the Farrell construction on a simplex code $\mathcal{S}_s$, when using a simplex code $\mathcal{S}_{s-2}$ as an anticode, as follows. By the construction of $\mathcal{C}^s$, the columns of the generator matrix for $\mathcal{C}^s$ are formed by all the vectors in $\mathbb{F}_2^s \setminus \{00v : v \in \mathbb{F}_2^{s-2}\}$. Then the columns of the generator matrix for the anticode are formed by the vectors in $\{00v : v \in \mathbb{F}_2^{s-2}\}$. This anticode has length $2^{s-2} - 1$ and maximum weight 2^{s-3}.*

IV. CONCLUSION

We presented a construction for four families of binary linear optimal LRCs with locality $r = 2$ and $r = 3$ (see Tab. I for some numerical examples). This construction is based on various anticodes. Besides the optimality with respect to the Cadambe–Mazumdar bound for a given locality, several of our families of codes fulfill the Griesmer bound with equality.

TABLE I
OPTIMAL BINARY LRC WITH LOCALITY TWO AND THREE.

$[n, k, d]$	Locality r	Reference	
$[28, 5, 14]$	2	Thm. 4	$m = 5, s = 3$
$[25, 5, 12]$	2		$m = 5, s = 4$
$[21, 5, 10]$	2		$m = 5, s = 5$
$[60, 6, 30]$	2	Thm. 4	$m = 6, s = 3$
$[57, 6, 28]$	2		$m = 6, s = 4$
$[53, 6, 26]$	2		$m = 6, s = 5$
$[21, 5, 10]$	2	Thm. 5	$m = 5, t = 4$
$[38, 6, 18]$	2		$m = 6, t = 5$
$[31, 6, 15]$	3	Thm. 6	$m = 6$
$[63, 7, 31]$	3		$m = 7$
$[24, 5, 12]$	2	Thm. 7	$s = 5$
$[48, 6, 24]$	2		$s = 6$

REFERENCES

[1] V. Cadambe and A. Mazumdar, "An upper bound on the size of locally recoverable codes," in *IEEE Int. Symp. Network Coding (NetCod)*, Jun. 2013, pp. 1–5.

[2] A. Dimakis, P. Godfrey, Y. Wu, M. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4539–4551, Sep. 2010.

[3] A. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, "A survey on network codes for distributed storage," *Proceedings of the IEEE*, vol. 99, no. 3, pp. 476–489, Mar. 2011.

[4] P. Farrell, "Linear binary anticodes," *Electron. Lett.*, vol. 6, no. 13, pp. 419–421, Jun. 1970.

[5] P. Gopalan, C. Huang, B. Jenkins, and S. Yekhanin, "Explicit maximally recoverable codes with locality," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5245–5256, Sep. 2014.

[6] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6925–6934, Nov. 2012.

[7] S. Goparaju and R. Calderbank, "Binary cyclic codes that are locally repairable," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2014, pp. 676–680.

[8] G. Kamath, N. Prakash, V. Lalitha, and P. Kumar, "Codes with local regeneration and erasure correction," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4637–4660, Aug. 2014.

[9] G. Kamath, N. Silberstein, N. Prakash, A. Rawat, V. Lalitha, O. Koyluoglu, P. Kumar, and S. Vishwanath, "Explicit MBR all-symbol locality codes," in *IEEE Int. Symp. Inf. (ISIT)*, Jul. 2013, pp. 504–508.

[10] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. North Holland Publishing Co., Jun. 1988.

[11] L. Pamies-Juarez, H. Hollmann, and F. Oggier, "Locally repairable codes with multiple repair alternatives," in *IEEE Int. Symp. Inf. (ISIT)*, Jul. 2013, pp. 892–896.

[12] D. S. Papailiopoulos and A. G. Dimakis, "Locally repairable codes," in *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 5843–5855, Oct. 2014.

[13] N. Prakash, G. M. Kamath, V. Lalitha, and P. V. Kumar, "Optimal linear codes with a local-error-correction property," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2012, pp. 2776–2780.

[14] N. Prakash, V. Lalitha, and P. Kumar, "Codes with locality for two erasures," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2014, pp. 1962–1966.

[15] A. Rawat, O. Koyluoglu, N. Silberstein, and S. Vishwanath, "Optimal locally repairable and secure codes for distributed storage systems," in *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 212–236, Jan. 2014.

[16] A. Rawat, D. Papailiopoulos, A. Dimakis, and S. Vishwanath, "Locality and availability in distributed storage," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2014, pp. 681–685.

[17] N. Silberstein and T. Etzion, "Codes and designs related to lifted MRD codes," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2011, pp. 2288–2292.

[18] N. Silberstein, A. Rawat, O. Koyluoglu, and S. Vishwanath, "Optimal locally repairable codes via rank-metric codes," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2013, pp. 1819–1823.

[19] M. Shahabinejad, M. Khabbazian, and M. Ardakani, "An efficient binary locally repairable code for hadoop distributed file system," in *IEEE Comm. Lett.*, vol. 18, no. 8, pp. 1287–1290, Aug. 2014.

[20] I. Tamo and A. Barg, "A family of optimal locally recoverable codes," in *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4661–4676, Aug. 2014.

[21] T. Westerback, T. Ernvall, and C. Hollanti, "Almost affine locally repairable codes and matroid theory," in *IEEE Inf. Theory Workshop (ITW)*, Nov. 2014, pp. 621–625.

[22] A. Zeh and E. Yaakobi, "Optimal linear and cyclic locally repairable codes over small fields," to appear in *IEEE Inf. Theory Workshop (ITW)*, Jerusalem, Israel, Apr. 2015.